AD_____

Award Number: W81XWH-08-1-0402

TITLE: Exploring the Pathogenic and Therapeutic Implications of
        Aberrant Splicing in Breast Cancer

PRINCIPAL INVESTIGATOR: William D Foulkes, Ph.D.

CONTRACTING ORGANIZATION: Jewish General Hospital
                          Montreal, H3T 1E2

REPORT DATE: July 2009

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
               Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: (Check one)

     x Approved for public release; distribution unlimited

REPORT DOCUMENTATION PAGE

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
31-07-2009	Annual	01-07-2008 to 30-06-2009

4. TITLE AND SUBTITLE

Exploring the Pathogenic and Therapeutic Implications of Aberrant Splicing in Breast Cancer

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-08-1-0402

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Foulkes, William D.
Email: william.foulkes@mcgill.ca

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Jewish General Hospital
Montreal, QC

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command, Fort Detrick, MD, 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

In this proposal, we set out to a) systematically monitor splicing variant profiles in breast cancer susceptibility genes and b) explore the role of alternative splicing in breast chemotherapy using a global strategy. In doing so, we hope to identify and validate candidate splicing variants involved in tumorigenesis using polony digital exon-profiling and functional assays. We are moving forward on four fronts – 1) the barcode methodology is in development; 2) we are working with state-of-the art capture arrays; 3) we are using the very latest RNA sequencing technology and 4) we are conducting comprehensive analyses of existing splice site variants in known breast cancer susceptibility genes. In part 3, we have generated over 5 million reads that have been aligned to the splice junction libraries, and we are using these reads to quantify and characterize alternative splicing events. Moreover, we will add value to this project by attempting to identify fusion proteins. In part 4, we have determined that the BRCA2 isoform known as BRCA2Δex12 is not associated with a recognizable phenotype. This work is now in press in Human Mutation.

15. SUBJECT TERMS

Breast cancer, splicing, single molecule analysis, RNA sequencing, BRCA2, Variants of Unknown Significance, Functional assays

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE		47	19b. TELEPHONE NUMBER (Include area code)

Table of Contents

Exploring the Pathogenic and Therapeutic Implications of Aberrant Splicing in Breast Cancer

W81XWH-08-1-0402

PRINCIPAL INVESTIGATORS: William D Foulkes MB PhD, Jun Zhu PhD

CO-INVESTIGATOR: Jacek Majewski, PhD

Introduction

In this proposal, we aimed to a) systematically monitor splicing variant profiles in breast cancer susceptibility genes and b) explore the role of alternative splicing in breast chemotherapy using a global strategy. In doing so, we hoped to identify and validate candidate splicing variants involved in tumorigenesis using polony digital exon-profiling and functional assays. This annual report summarizes our progress in the first year of this two-year award.

Body

Verbatim – from S.O.W

Objective 1: To profile splicing patterns of 100 breast cancer-related genes in 30 normal and tumor breast

cell lines

Objective 2: To profile splicing patterns of 100 cancer-related genes in lymphocytes and breast tissues

Objective 3: Data analysis and validation to determine the role of aberrant splicing in breast tumourgenesis *(months 8-12)*

These objectives can be summarized in terms of the expected Results/Deliverables from Year 1: 1) Optimizing the method of bar-code PCR-sequencing for studying alternative splicing at candidate loci. 2) Identifying splicing variants associated with increased breast cancer susceptibility derived from the candidate loci. Year 2 focuses more on drug sensitivity, and how splicing patterns may determine the response to certain chemotherapeutic agents.

<u>1) Optimizing the method of bar-code PCR-sequencing for studying alternative splicing at candidate loci</u>

As the key subject of this synergistic IDEA award was the development of a molecular barcode, as a prelude to large scale analysis of splicing patterns present in the breast cancer genome, we undertook a the first step to evaluate the "Molecule Barcode" technology to monitor alternative splicing in tumor samples, we have performed a series of proof-of-concept experiment. The results showed that the effective sequencing length can be virtually extended from 36 bp to 300-400 bp. While the results are encouraging, we observed that the coverage is biased towards the middle of reconstructed molecules. In parallel with attempts to improve our technology, and to better monitor splicing variants in tumor samples, we switched to an emerging alternative strategy, NimbleGen Sequence-Capture array that can selectively enrich genes of interest (3-20 Mbps) for deep sequencing analysis. The captured fragments were then subjected to 454

pyrosequencing to obtain a longer read length (~400bp), which enables better identification of alternative splicing and mutation detection.

We had worked with NimbleGen to design a custom capture array for detail characterizing a larger set breast cancer related genes. The original design was based on 385K array (Nov, 2008) and training date was set for early 2009. Both Ting Ni (post-doctoral fellow, funded by this award, based at Duke University) and Lili Li (PhD candidate, funded from this award, based at McGill University) took the training courses using supplied 385K capture array by NimbleGen during April 7 - 9, 2009. Average 288-345 fold of enrichment using supplied genomic DNA suggested the success of training. We then use the real samples of the breast cancer cell line BT20 and drug-resistant BT20 cell lines to practice by ourselves. Real-time PCR results suggested substantial enrichment of desired regions (average enrichment of 267 fold).

During our interactions with NimbleGen, the 2.1M capture array capture array became commercially available with an early release in Feb, 2009. We were offered by NimbleGen to have a free upgrade from 385K array to the most advanced 2.1M array. Given that the new custom array has approximately 6-fold higher capacity, we decided to broaden our research using the high-density array. The new custom array was design in early March, which before the actual training was taken place. After the training, we requested the 2.1M custom array to be delivered. However, there was a logistic mistake from Nimblegen and our order remained showing as 385K array in their ordering system. The issue had been resolved in June, 2009. The array order is currently in production department and we expect to have the array delivered within next 2-4 weeks.

While waiting for the custom array to be delivered, we have constructed two libraries with BT20 cell line and its derivative that resistant to dastanib, an important new tyrosine kinase drug that has considerable activity in leukemia and in some solid tumors, including breast cancer. We have worked out an improved RNA-Seq protocol to keep the strand information during library construction. Therefore, the library will allow us not only to detect splicing variants but also antisense transcripts in cell transcriptome. We have obtained initial information that antisense transcripts are broad phenomena of eukaryotic transcriptomes and may play essential roles in mammalian gene regulation. The innovative strategy we developed, in fact, can reliably capture this hidden layer of mammalian transcriptomes and is expected to provide novel insights into altered gene expression in addition to alternative splicing.

Lastly, it is worth pointing out the NimbleGen capture array is directly coupled with 454 pyrosequencing platform. To prepare further study with large number of samples, we also tested the strategy to convert the captured sequences to library amenable for Illumina/Solexa Sequencing (which is approximately 10 times more cost-effective than 454 pyrosequencing). This led to a optimized protocol by which the 454 library can be adapted for Solexa sequencing through a circularization step.

In addition to the analysis of the BT20 lines, In addition, through collaboration with the Breakthrough Breast Cancer Institute in the UK, we have collected RNA samples from 34 breast tumor lines that represent 5 different types of breast cancers classified by expression profiles: luminal epithelial-like subtypes A and B, basal epithelial-like, ERBB2-amplification associated and normal breast-like subtypes. The fact that the response profiles of these lines to Dasatinib have been published, will allow us to assess whether or not the genetic alterations identified in a induced resistance model can be applied to intrinsic resistance. Thus, as soon as our pipeline is well-developed we will be able to test and validate candidate genes using this well-characterised set of breast cancer cell lines.

A third strategy we have used is to sequence the RNA of a breast cancer cell line directly. Because the pace of technological innovation is substantially faster than the granting process, our original proposal to analyze the splicing patterns of 100 candidate breast cancer genes (using the barcode and/or capture array strategy) has been augmented this approach with state-of-the art whole transcriptome sequencing. In other words, in this part of the project, we aim to analyze not the 100 "top" candidate breast cancer susceptibility genes, but instead all ~24,000 genes. This project has been led by PI Foulkes and co-I Majewski. To date, we have used Illumina/Solexa RNA sequencing to generate over 40 million, 50 bp reads from the transcriptome of the well known breast cancer cell line, HCC1937. The reads were mapped to the transcriptome using ELAND and CASSAVA software, and visualized using the Illumina Genome Studio. Summary of the mapping is presented in Table 1 below.

Table 1 Summary of Illumina Genome Analyzer Sequencing Reads

Category	Total Number of Reads
Non-Mapping	1,515,068
Repeat Masked	8,789,665
Too Many N's in Sequence	1,316
Aligned to Reference Genome	24,701,485
Aligned to Splice Junctions	5,297,332
Total	40,304,866

Over 5 million reads are aligned to the splice junction libraries, and we are using these reads to quantify and characterize alternative splicing events. However, one of our main interests is to identify novel splice isoforms (not present in most reference splice libraries). To this end, we are pursuing three approaches: 1) Mapping sequencing reads to exhaustive splice junction libraries representing all possible splice junctions within each gene; 2) monitoring excess reads that map to normally intronic sequences – this is meant to detect novel intron retention events; 3) creating a library and heuristic algorithms for mapping reads to an exhaustive library joining all exons across all of the genes in the human genome – this last approach will allow us to detect all known and novel alternative splicing events, trans-splicing (across genes), as well gene fusion events that are known to occur in many tumors. We discuss this last, most inclusive approach below.

Gene fusions are the result of chromosomal alterations involving two genes. These chimeras may have severe phenotypic effects, such as the well-studied *BCR-ABL1* fusion protein implicated in chronic myelogenous leukemia (Shtivelman et al., 1985) and *TMPRSS2-ERG* found in many cases of prostate cancer (Tomlins et al., 2005). New efforts using high-throughput sequencing have resulted in new discoveries of gene fusions. This has prompted interest in determining whether these chromosomal aberrations may be specific to cancer and if they are, may theoretically serve as an ideal diagnostic and therapeutic target (Prensner and Chinnaiyan, 2009).

To detect gene fusions using our RNA sequencing data, we followed a similar approach described by Maher et al. (2009). Of a total of 40 million 50 bp reads sequenced from the HCC1937 cell line, approximately 1.5 million reads were determined to be non-mapping after analyzing the data through the Illumina pipeline (version 1.3). Non-mapping reads are those which did not have a unique alignment to the reference

genome. We limited our gene fusion analysis to only these reads as they may potentially characterize breakpoints of gene fusions not found in the reference genome.

Firstly, we downloaded from the UCSC Genome Database (hg18 assembly) the set of exon genomic sequences from all mature mRNA RefSeq transcripts. As the mRNA of gene fusions would typically involve the fusion of exon sequences from two different genes, we retained only the boundary sequences of each exon (i.e. 49 bp sequence from the left boundary of an exon and 49 bp sequence from the right boundary of an exon), in order to identify reads that may span such exon-exon boundaries. Blat (Kent, 2002) was used to align the non-mapping reads against this list of exon boundary sequences. Using a set of in-house scripts written in Python, we filtered the results for alignments such that a read was partially aligning to an exon boundary. Of these partial alignments, we further identified whether its remaining unaligned sequence aligned to another exon region either from the Blat analysis or by simple string matching techniques.

Figures 1 and 2. Genome Studio screenshots, one showing an example of an alternatively spliced gene (FOXM1) and one showing a non-expressed gene, PTEN, which is known to be deleted in this cell line.
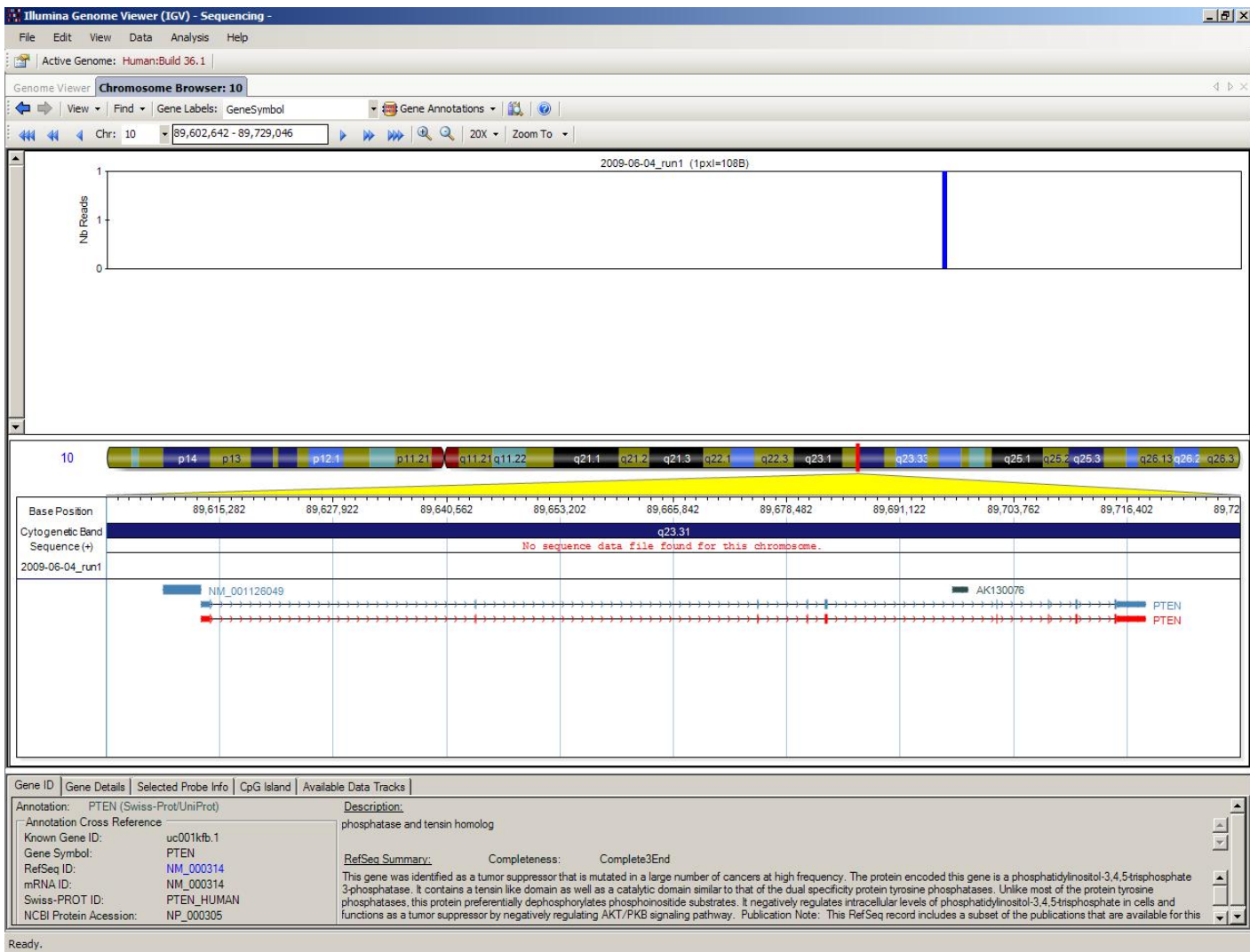

Figure 1 Screenshot from Genome Studio of PTEN, which is known to be not expressed in the HCC1937 cell line. Here, only one read is mapped to this region and is within intronic region of PTEN.

4

Illumina Genome Viewer (IGV) - Sequencing -

File   Edit   View   Data   Analysis   Help

Active Genome: Human:Build 36.1

Genome Viewer | Chromosome Browser: 12

View ▾ | Find ▾ | Gene Labels: GeneSymbol   ▾ | Gene Annotations ▾ |

Chr: 12 ▾ | 2,837,110 - 2,857,564 | | 10X ▾ | Zoom To ▾

2009-06-04_run1 (1pxl=17B)

316

158

0

Nb Reads

12   p13.31   p12.3   p12.1   q12   q13.13   q14.1   q15   q21.1   q21.2   q21.31   q21.33   q22   q23.1   q23.3   q24.31   q24.32   q24.3

Base Position   2,839,150   2,841,190   2,843,230   2,845,270   2,847,310   2,849,350   2,851,390   2,853,430   2,855,470   2,857

Cytogenetic Band   p13.33

Sequence (+)   No sequence data file found for this chromosome.

2009-06-04_run1

U90917   FOXM1   FOXM1   FOXM1   FOXM1   FOXM1   FOXM1

Gene ID | Gene Details | Selected Probe Info | CpG Island | Available Data Tracks |

Annotation: Select a Gene...   Description:

Annotation Cross Reference

Known Gene ID:

Gene Symbol: No Data

RefSeq ID: RefSeq Summary: Completeness:

mRNA ID:

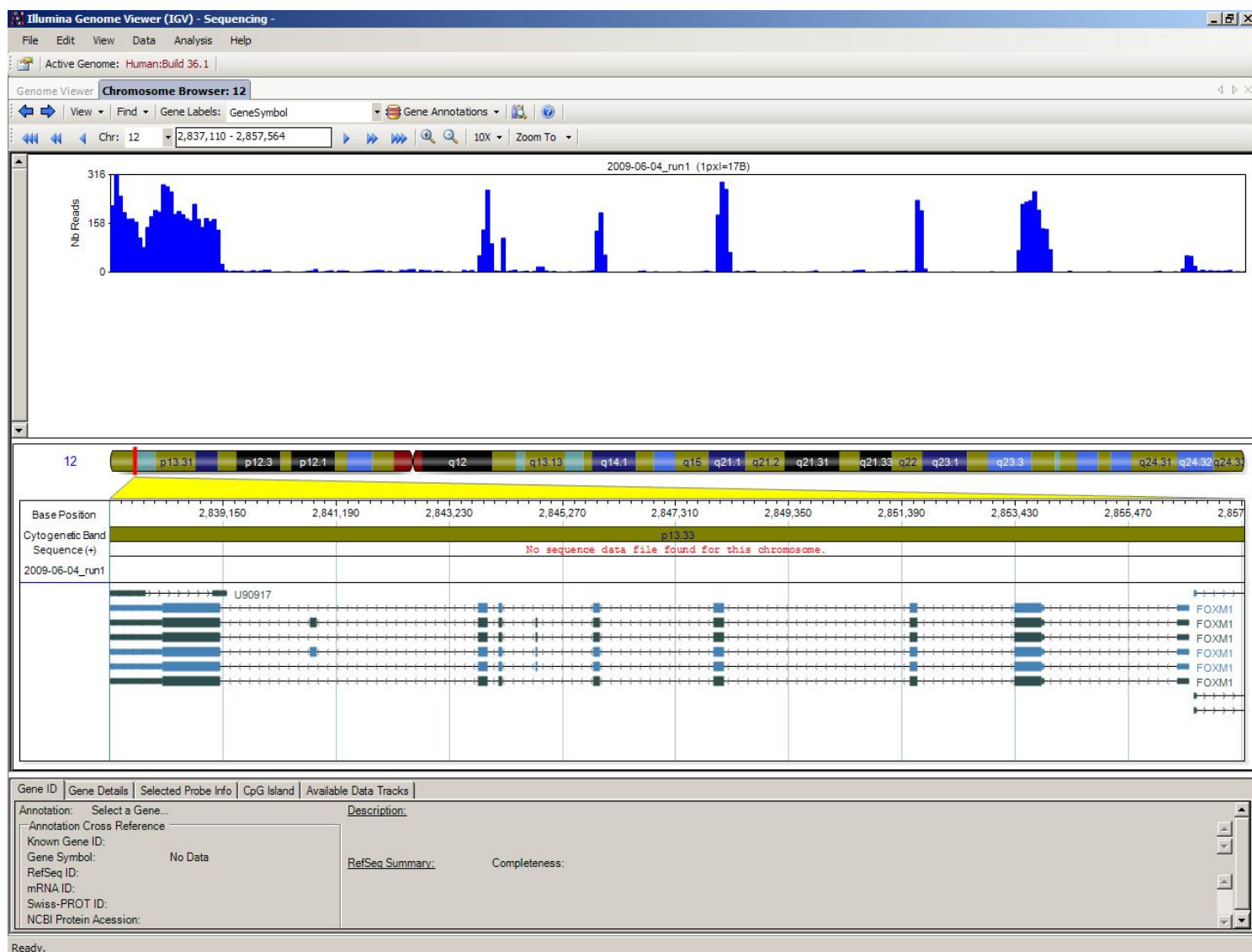Swiss-PROT ID:

NCBI Protein Acession:

Ready.

Figure 2 Example of an alternatively spliced gene. Here, FOXM1 contains two isoforms, one of which has two cassette exons spliced out. Based on the reads summary, it can be seen that the shorter isoform is being represented.

In addition to gene fusion detection, our approach will lead to the detection of novel splicing events and transcript isoforms. It will be possible to observe previously uncharacterized exon-exon junctions from a particular gene, which may represent a transcript isoform.

Finally, as part of this project, we have focused on splice variants of one particularly important breast cancer susceptibility gene, BRCA2. Previous studies have suggested that the variant NM_000059.3 c.6853A>G, BRCA2: p.I2285V (also known as c.7081A>G) is unlikely to be a deleterious allele and in support of this, we observed that BRCA2: p.I2285V co-occurs in *trans* with *BRCA2*: c.5946delT (more commonly known as 6174delT). However, we also noted that when compared with wild-type BRCA2, BRCA2: p.I2285V resulted in increased exclusion of exon 12, leading to the accumulation of a shortened form of BRCA2 protein. The functional significance of this shortened isoform of BRCA2 is uncertain. Allelic complementation in *Brca2* null mouse embryonic stem cells was performed by our collaborator Shyam Sharan (in his letter of collaboration in the grant submission he wrote *"We are currently working on several other variants, including I2285V, which although not a high-risk allele, is associated with increased exclusion of exon 12.*

5

Using our system, we will be able to test splice variants that you discover".) He showed that both p.I2285V and an allele completely lacking exon 12 (BRCA2Δex12) were functionally indistinguishable from wild-type BRCA2. This was based on their sensitivity to DNA-damaging agents, effect on irradiation-induced Rad51 foci formation, homologous recombination and overall genomic integrity. Taken together, our clinical and functional studies strongly suggested that exon 12 is functionally redundant and therefore missense variants in this exon are likely to be neutral. We believe that such comprehensive functional studies will be important adjuncts to genetic studies of variants.

Thus, although the original the molecular barcode strategy has not yet been perfected, we are moving forward on four fronts – as stated, 1) the barcode methodology is in development; 2) we are working with state-of-the art capture arrays; 3) we are using the very latest RNA sequencing technology and 4) We are conducting comprehensive analyses of existing splice site variants in known breast cancer susceptibility genes. As stated above, in part 3, we have generated over 5 million reads that have been aligned to the splice junction libraries, and we are using these reads to quantify and characterize alternative splicing events. Moreover, we will add value to this project by attempting to identify fusion proteins, as discussed above. In part 4, we have determined that the BRCA2 isoform known as BRCA2Δex12 is not associated with a recognizable phenotype.

Key Research Accomplishments

- Publication in press from the Foulkes group
- Development of novel approaches to analyzing RNA Seq data
- Substantial enrichment of desired regions on the Nimblegen array (average enrichment of 267-fold) has been confirmed by real-time PCR

Reportable Outcomes

- Manuscript partly funded by Department of Defense: Lili Li, Kajal Biswas, Laura Anne Habib, Sergey G. Kuznetsov, Nancy Hamel, Tomas Kirchhoff, Nora Wong, Susan Armel, George Chong, Steven A. Narod, Kathleen Claes, Kenneth Offit, Mark E Robson, Stacey Stauffer, Shyam K. Sharan and William D. Foulkes: Functional redundancy of exon 12 of *BRCA2* revealed by a comprehensive analysis of the *BRCA2* variant c.6853A>G, p.I2285V. *Human Mutation*, in press. – this ms is concerned with a breast cancer susceptibility gene, BRCA2. An isoform of BRCA2 known as BRCA2Δex12 exists. We have explored the functional significance of lacking exon 12 of BRCA2.
- #1 Seminar (Dr Zhu): "NextGen Sequencing analysis of eukaryotic transcriptomes", NHLBI (July, 2009)
- #2 Seminar (Dr. Zhu): "Single-molecule barcoding technology for RNA-Seq", Craig Venter Institute (April, 2009)
- #1 Meeting Presentation (Dr. Zhu): "Strand-specific transcriptome sequencing for discovery of natural antisense transcripts" Ting Ni, Kang Tu, Shen Song, Zhong Wang, Jun Zhu, RNA 2009: the Fourteenth annual meeting of the RNA Society (May, 2009)
- #2 Meeting Presentation (Lili Li) : "The Clinical Importance of Weak Exons in *BRCA2*?" This was a question originating from the comprehensive analysis of I2285V in Exon12 - Annual report for the Training Program of Innovative Technologies In Multidisciplinary Health Research University of Manitoba, Winnipeg, MB, Canada (March 23th, 2009)
- #3 Meeting Presentation (Lili Li): "The Clinical Importance of a "Weak" Exon in *BRCA2*? Human Genetics Graduate Research Day, McGill University, Montreal, QC, Canada (May29[th], 2009)

- Patent application: Zhu J., Ni T. and Gao Y. (2008) Barcode multiplex PCR and single-molecule decomposition. (US patent 61/131,279).
- #1 Grants awarded: Zhu, Foulkes and Majewski awarded Susan G Komen for the Cure grant (2008).
- #2 Grants awarded: Foulkes and Majewski awarded Canadian Breast Cancer Research Alliance IDEA grant to discover novel fusion proteins in BRCA1-related breast cancer (2009)
- #3 Grants awarded: Zhu awarded a NSF grant entitled "Genome-wide exploring miRNA Network motifs" NSF-0822033 with a recent "creative extension" 9/1/2008-8/31/2012.
- #1 Grants applied for: Majewski and Foulkes have applied for a joint Quebec-China informatics grant
- #2 Grants applied for: Zhu has have applied for a grant entitled: "Functional elements in miRNA processing", NHGRI, R21
- Training: A Masters student will join the McGill group in September 2009.

Conclusion

In this proposal, we have set out to evaluate the importance of splicing for breast cancer biology. We are employing a number of state-of-the-art technologies – in particular, we are developing the very novel "Molecule Barcode" technology. As this has been a little slower than we would have liked, in addition we have followed two other avenues – a custom-made capture array to study splicing events in breast cancer and RNA sequencing, which is a relatively unbiased approach to the problem. Finally, we have carefully evaluated the significance of one particular isoform of BRCA2, and have shown that it is not associated with an increased risk for breast cancer. This is important for the families whose members carry this variant. In the future, we will continue to focus on developing our innovative technology development but will also pursue alternative strategies, as described above. If necessary, we will also continue to collaborate with Dr. Shyam Sharan in attempting to classify splicing variants in the important breast cancer susceptibility genes, BRCA1 and BRCA2. The knowledge gained from this project thus far could directly benefited women with breast cancer and those at risk of breast cancer. We hope that our approach will lead to significant insights into the more general question of the importance of alternative splicing in breast cancer biology.

References

Kent, W. J. (2002). BLAT--the BLAST-like alignment tool. *Genome Res*, 12(4), 656-664.

Maher, C. A., Kumar-Sinha, C., Cao, X., Kalyana-Sundaram, S., Han, B., Jing, X., et al. (2009). Transcriptome sequencing to detect gene fusions in cancer. *Nature*, 458(7234), 97-101.

Prensner, J. R., & Chinnaiyan, A. M. (2009). Oncogenic gene fusions in epithelial carcinomas. *Curr Opin Genet Dev*, 19(1), 82-91.

Shtivelman, E., Lifshitz, B., Gale, R. P., & Canaani, E. (1985). Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. *Nature*, 315(6020), 550-554.

Tomlins, S. A., Rhodes, D. R., Perner, S., Dhanasekaran, S. M., Mehra, R., Sun, X., et al. (2005). Recurrent Fusion of TMPRSS2 and ETS Transcription Factor Genes in Prostate Cancer. *Science*, 310(5748), 644-648.

Appendices

See attached ms: Li et al, Human Mutation, in press, 2009

Human Mutation

WILEY

Functional redundancy of exon 12 of BRCA2 revealed by a comprehensive analysis of the BRCA2 variant c.6853A>G, p.I2285V

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Functional redundancy of exon 12 of *BRCA2* revealed by a comprehensive analysis of the *BRCA2* variant c.6853A>G, p.I2285V

Lili Li*[1], Kajal Biswas*[2], Laura Anne Habib[1], Sergey G. Kuznetsov[2], Nancy Hamel[3], Tomas Kirchhoff[4], Nora Wong[5], Susan Armel[6], George Chong[7], Steven A. Narod[8], Kathleen Claes[9], Kenneth Offit[4], Mark E Robson[4], Stacey Stauffer[2], Shyam K. Sharan[2] William D. Foulkes[1,3,5]

[1]Program in Cancer Genetics, Departments of Oncology and Human Genetics, McGill University, Montreal, QC, Canada; [2]Mouse Cancer Genetics Program, Center for Cancer Research, National Cancer Institute at Frederick, Frederick, MD, USA; [3]Research Institute of the McGill University Health Centre; [4]Clinical Genetics Service, Memorial Sloan Kettering Cancer Centre, New York, NY USA; [5]Department of Medical Genetics and Cancer Prevention Centre, Sir Mortimer B. Davis-Jewish General Hospital, Montreal, QC, Canada; [6]Familial Breast and Ovarian Cancer Clinic, Princess Margaret Hospital, University of Toronto, ON, Canada; [7]Department of Pathology, Sir Mortimer B. Davis-Jewish General Hospital, Montreal, QC, Canada; [8]Women's College Hospital Research Institute, University of Toronto, ON, Canada; [9]Center for Medical Genetics, Ghent University Hospital, Belgium

*These authors have contributed equally

Corresponding author:
William D Foulkes
Room E-757
Cancer Prevention Centre
SMBD-Jewish General Hospital
3755 Côte Ste Catherine Road, Montreal, H3T 1E2, QC, Canada
Tel. 1 514 340 8222 x3965
Fax 1 514 340 8712
william.foulkes@mcgill.ca

Abstract

Variants of Unknown Significance (VUS) in *BRCA1* and *BRCA2* are common, and present

significant challenges for genetic counseling. We observed that *BRCA2*: c.6853A>G

(p.I2285V) (BIC name: 7081A>G) co-occurs in trans with the founder mutation c.5946delT

(p.S1982RfsX22) (BIC name: 6174delT), supporting the published classification of

p.I2285V as a neutral variant. However, we also noted that when compared with wild-type

BRCA2, p.I2285V resulted in increased exclusion of exon 12. Functional assay using

allelic complementation in *Brca2* null mouse embryonic stem cells revealed that p.I2285V,

an allele with exon 12 deleted and wild-type BRCA2 were all phenotypically

indistinguishable, as measured by sensitivity to DNA-damaging agents, effect on

irradiation-induced Rad51 foci formation, homologous recombination and overall genomic

integrity. An allele frequency study showed the p.I2285V variant was identified in 15/722

(2.1%) Ashkenazi Jewish cases and 10/475 (2.1%) ethnically-matched controls, odds ratio:

0.99 (95% confidence interval: 0.44-2.21), $P = 0.97$. Thus the p.I2285V variant is not

associated with an increased risk for breast cancer. Taken together, our clinical and

functional studies strongly suggest that exon 12 is functionally redundant and therefore

missense variants in this exon are likely to be neutral. Such comprehensive functional

studies will be important adjuncts to genetic studies of variants.

Key words: BRCA2, unclassified variants, co-occurrence, exon splicing enhancer, exon

skipping, in-frame deletion, neutral variant, Embryonic Stem (ES) cells

Introduction

Variants of unknown significance (VUS), also known as unclassified variants (UV),

account for a small, but nevertheless significant, proportion of all variants identified in

BRCA1 (OMIM#113705) and *BRCA2* (OMIM#600185). Interpreting such variants pose

significant challenges for both clinicians and patients. The unclassified variant

NM_000059.3: c.6853A>G (p.I2285V) appears to be restricted to the Ashkenazi Jewish

population and has been reported 79 times in the BIC database (http://research.nhgri.

nih.gov/bic/) (BIC name: 7081A>G) [Mutation Nomenclature used in this paper:

nucleotide numbering reflects cDNA numbering, with +1 as the A of the ATG initiation

codon in the reference sequence (BRCA2: NM_000059.3) (www.hgvs.org/mutnomen). The

initiation codon is codon 1. Traditional numbering based on the BIC mutation database is

also provided (http://research.nhgri.nih.gov/bic/)]. It is the most frequently reported variant

in exon 12, and was identified 97 times in 68,000 tests performed by Myriad Genetics

Laboratories. Several studies of *BRCA1/2* mutations in women with breast or ovarian

cancer have identified p.I2285V in affected women and have designated p.I2285V as a

VUS (Shih et al., 2000; Pal et al., 2005). We were requested to offer genetic counseling to

an unaffected member of a family who carries this variant. In this context, it was reported

to us (by a Canadian clinical molecular genetics laboratory) to be possibly associated with

aberrant splicing of exon 12 of BRCA2, resulting in a truncated protein. If exclusion of

exon 12 occurs, the result is an in-frame deletion of 32 amino acids (codon 2281-2312).

This isoform is known as BRCA2ΔE12.

Splice site mutations are a common cause of exclusion of an exon and many have been

identified in *BRCA1* and *BRCA2* (Frank et al., 2002). More recently, variants that do not

alter the consensus splice site sequences but affect splicing have been identified (Liu et al.,

2001). These mutations result in the interruption of exon splicing regulators and have been

increasingly recognized as important causes of exon exclusion or intron inclusion (Cartegni

et al., 2002). Since the BRCA2: p.I2285V was reported to be associated with increased

exclusion of exon 12 in one family, we were interested to establish if this were the case in

other individuals carrying this variant, and whether disruption of an ESE was a likely

mechanism. Moreover, we wondered if exclusion of exon 12 was associated with an

identifiable phenotype.

Interestingly, it had been previously reported that breast cancer tissues show increased

amount of E12 skipping transcript (BRCA2ΔE12) (Bieche and Lidereau, 1999), but no

specific function has been attributed to this polypeptide. Notably, several cell lines such as

MDA-MB-468, MCF-7, OVCA420, DOV13 and SKOV3 (among others) have increased

expression of this variant. In addition, the prostate cancer cell line DU-145 expresses

BRCA2 ΔE12 at much greater levels than the full-length isoform (Rauh-Adelmann et al.,

2000). Despite these findings, co-segregation and family history analyses have shown that

p.I2285V is unlikely to be a highly penetrant allele (Easton et al., 2007). This view is

supported by the recently published work of Tavtigian and colleagues, which predicts that

p.I2285V is a neutral variant, because the two independent measures used (missense

mutations outside of the DNA binding domain and family history likelihood ratio) place

this variant in a group that has a prior probability of <0.05 (Tavtigian et al., 2008). The

potential effect of a variant on splicing, however, is not measured in these analyses. Taking all the observations together, it seems possible that this variant affects splicing and thus has a subtle effect on risk, perhaps akin to the risk associated with so-called low penetrance breast cancer genes identified recently (Pharoah et al., 2008). Here, we describe investigation of the relationship between this variant and exclusion of exon 12; b) the functional consequences of exclusion of exon 12 of *BRCA2* and c) the potential effect of this variant on breast cancer risk.

Materials and Methods

BRCA2: p.I2285V pedigrees

We collected information from four families where at least one individual carried the BRCA2: p.I2285V variant. These pedigrees are shown in supplementary figures 1a-d. Lymphocyte RNA and DNA was extracted from one carrier of the variant from each family and were used in the experiments described below. All participants signed an informed consent form. The details of each family are described in the figure legend accompanying supplementary figure 1.

Testing for inclusion or exclusion of exon 12 in humans

We confirmed the co-occurrence of *c.6853A>G* / p.I2285V (BIC name: *7081A>G*) with another deleterious BRCA2 mutation *c.5946delT* / p.S982RfsX20 (BIC name: *6174delT*), using a blood sample from a patient carrying both alleles (see supplementary figure 2 and accompanying legend). We used RNA extracted from blood samples from this proband and three others (supplementary figure 1) using PAXgene (Qiagen) kit and cDNA was

synthesized using Sensiscript (Qiagen). We used the following PCR primers to semi-

quantitatively compare the degree of exon 12 skipping between the cells lines the wild-type

line (referred to as WLT) and the double variant line (referred to as DBV): 1) for exon 12

inclusive transcript: Forward (located in exon 12) : 5' – CCCTTATCTTAGTGGG

AGAACCCTCA-3', Reverse (located in exon 13): 5'-

AGTTGTGCGAAAGGGTACACAGGT-3'; 2) for constitutive exon 11: Forward (located

in exon 11): 5' –TTGAGGTAGGGCCACCT GCATTTA-3', Reverse (located in exon 11):

5'-ATCCAATGCCTCGTAACAACCTGC-3'; 3) for exon 12 skipping transcript: Forward

(located at junction of exon 11 and exon 13): 5'-ATCTTAGTGGGCACAATAAAAG-3';

Reverse (located in exon 13): 5'-CAGAAATTCTTGA CCAGGTGCGGTA-3'. The relative

intensity of the PCR products was measured by using GeneTool software (Syngene).

Testing for the functional significance of exclusion of BRCA2 exon 12 in mouse

embryonic stem (ES) cells

To test the functional significance of p.I2285V and BRCA2ΔE12 variants, mutant BAC

DNAs were generated in the human *BRCA2* BAC RP11-777I19 by recombineering. A

mutant BAC DNA containing the p.I2285V variant was constructed by two step 'hit and

fix' method as described previously (Yang and Sharan, 2003). The BRCA2ΔE12 variant

was generated in the BAC by using *galK* selection and counter-selection method as

previously described (Warming et al., 2005) The sequences of the oligonucleotides are

available upon request. Mutant BAC DNAs were then introduced in PL2F7 mouse ES cells

followed by deletion of the endogenous mouse *Brca2* gene by Cre-mediated recombination

(Kuznetsov et al., 2008). Recombinant clones lacking endogenous *Brca2* were selected in

HAT media as described previously (Kuznetsov et al., 2008). We tested the efficiency of

the DNA repair function by challenging the recombinant cells with DNA-damaging agents:

N-methyl-N'-nito-N-Nitrosoguanidine (MNNG), mitomycin C (MMC), methylmethane

sulfonate (MMS), cisplatin, ultra-violet light and γ-irradiation. Significance of the survival

difference between mutant and wild-type cells was assessed by two-tailed t-test at drug

concentrations corresponding to LD_{50}. We tested the efficiency of homologous

recombination, RAD51 foci formation and karyotype of the wild-type and the recombinant

cells as described previously (Kuznetsov et al., 2008).

For the RT-PCR, we synthesized cDNA using RNA isolated from mouse ES cells

expressing wild-type, I2285V and Δexon12 mutant human *BRCA2* gene using SuperScript

III reverse transcriptase (Invitrogen). Region flanking exon 12 was amplified with a

forward primer B2ex11FRT (5'-CCAAGTCATGCCACACATTC-3') and a reverse primer

B2ex14RRT (5'-ATTCTTGACCAGGTGCGGTA-3') from exons 11 and 14, respectively,

and the PCR conditions included: one cycle of 94°C for 5 min, 35 cycles of 94°C for 30

sec, 55°C for 30 sec, 72°C for 30 sec, followed by incubation at 72°C for 5 min.

Amplification products were separated in a 1.6% agarose gel.

Allelic association study

As controls, the study enrolled 475 healthy women collected through the New York Cancer

Project (NYCP). The NYCP is a cohort study involving consent for biospecimen collection

and follow-up of 8,000 healthy volunteers in the same geographical region as the cases

used in this study (Mitchell et al., 2004). For the cases, we genotyped 722 consecutive

breast cancer cases unselected for family history or BRCA1/2 status seen at Memorial

Sloan-Kettering Cancer Center and collected for the study under IRB approved protocols.

All individuals in the study were self identified as Ashkenazi Jewish. For genotyping of

BRCA2 I2285V we used TaqMan allelic discrimination assay under standard

manufacture's protocol (Applied Biosystems). The alleles were called using SDS 2.1

software (Applied Biosystems).

Results

Conservation and co-occurrence

When we examined the amino acid sequence at position p.2285 of human *BRCA2* with

Align-GVGD (http://agvgd.iarc.fr/alignments/) (Tavtigian et al., 2006; Mathe et al., 2006),

we found it to be relatively conserved across vertebrate *BRCA2* orthologs. This residue is

located within a conserved motif encoded by exon 12. From humans through to pufferfish

(tetraodon), this position is either Ile or Leu. Val at this position can be considered to be a

relatively conservative substitution because the three amino acids are structurally similar,

each with a nonpolar side chain.

We confirmed that *c.6853A>G* co-occurs with a deleterious variant by testing 85

Ashkenazi Jewish individuals from Montreal and Toronto who had previously been found

to carry the *6174delT* (*c.5946delT*) variant in *BRCA2*. We identified one living person who

carries both variants (supplementary figure 1). We re-contacted this person and established

a lymphoblastoid cell line (LCL), hereon referred to as DBV (for double variant). Being

unable to study other family members, we could not establish phase genetically. Instead, by

using RNA from this cell line, we showed that exon 12 inclusion and exclusion transcripts

were produced from both alleles, but the degree of exclusion was much greater with for the

I2285V-associated allele than for the *c.6174delT*-associated allele. We were able to

demonstrate that these variants occur in *trans* (supplementary Figure 2). This supports the

previous segregation and family history data showing that p.I2285V is very unlikely to be a

highly penetrant allele (as discussed above). Nevertheless, we were left with the question of

the splicing out of exon 12 – is it frequently associated with BRCA2:p.I2285V, and if so,

what is the consequence of this?

Effect of I2285V on the BRCA2 mRNA

The clinical protein truncation test (PTT) of *BRCA2* on one individual carrying this variant

indicated that exon 12 skipping occurs to a greater degree in this individual than the level

in the control samples (data not shown). We searched for splicing regulatory elements

located in exon 12 by using the following on-line prediction tools: ESEFinder:

http://rulai.cshl.edu/tools/ESE; RESCUE-ESE: http://genes.mit.edu/burgelab/rescue-ese/

ESRsearch:http://ast.bioinfo.tau.ac.il/ESR.htm and

PESX: http://cubweb.biology.columbia.edu/pesx/. Using RESCUE-ESE, we identified

putative ESEs at the region where the *BRCA2: c.6853A>G* variant is located. Notably, the

variant "G" allele disrupts three of overlapping putative ESEs (Figure 1A). By studying

exon 12 retention in 4 individuals who carry the p.I2285V variant and two controls, we

were able to show quantitatively that this variant significantly increases the exclusion of

exon 12 of *BRCA2* (Figure 1B). The test we used here was to see if the variant was retained

in exon 12 by capillary sequencing of exon 12 inclusion transcripts, as a variant promoting

exon 12 skipping would result in a diminished signal of itself in the retained exon 12. As

shown in Figure 1B, among 3 heterozygotes (I2285V-a, I2285V-b and I2285V-c), the

c.6853A>G variant (resulting in p.I2285V) is barely detectable in retained exon 12,

indicating the "G" variant strongly promotes exon 12 skipping. This causes extremely

skewed A:G allele ratio in retained exon 12, to a degree that the G allele (C as sequenced

with a reverse primer) is below the detection power of capillary sequencing. The

interpretation of the results is somewhat more complicated for the lymphoblastoid cell line

DBV (Figure 1B, fifth sample), as this cell line is derived from individual compound

heterozygote for *c.6853A>G* and *c.6174delT* (discussed above and shown in supplementary

figure 2). It is possible that the ratio of A:G has been reverted by the presence of a degree

of nonsense-mediated decay of the A allele intact transcript, resulting from the upstream

c.6174delT mutation on this allele, which introduces a premature stop codon. Despite the

fact that the upstream *c. 6174delT* may equally causes some decay of the *BRCA2ΔE12*

transcript produced from this allele in individual DBV (in which *c.6853A>G* and

c.6174delT are present in *trans*), a greater degree of exon 12 skipping was observed in this

cell line (as the sum of the skipping products from both alleles) than in the wild-type

counterpart lymphoblastoid cell line WLT (Figure 1C), suggesting that *c.6853A>G*

enhances exon 12 skipping. Therefore, the biologically relevant outcome of this single

base substitution is likely to be the disruption of a putative ESE, resulting in increased

exclusion of exon 12. The creation of a novel amino acid could also have some biological

significance, but this seems less likely as the change of I to V is only just outside the

tolerated range. Our next goal was to establish if this splicing out of exon 12 has any

functional consequences.

Functional test of c.6853A>G variant and complete deletion of exon 12

To test the functional significance of deletion of exon 12, mouse ES cell based functional

assay was used. The *BRCA2ΔE12* variant was generated in the BAC DNA coding for

human BRCA2 and introduced into mouse ES cells containing a single functional copy of

Brca2 that can be deleted by Cre-mediated recombination. Mouse ES cell are not viable

without a functional copy of *Brca2*. After Cre-mediated deletion of endogenous mouse

Brca2, we tested the ability of human transgene to rescue the lethality of *Brca2*-null ES

cells. We tested 6 independent ES cell clones expressing the *BRCA2ΔE12* transgene and 5

clones expressing *c.6853A>G* variant (resulting in p.I2285V) for viability and found no

difference compared to the wild-type (data not shown). RT-PCR analysis of mouse ES cells

expressing *BRCA2ΔE12* variant confirmed the deletion of 96 bp encoded by exon 12

(Figure 2, lanes 5 and 6). These cells do not express the wild-type transcript. Similarly, ES

cells expressing the p.I2285V variant resulted predominantly in transcripts lacking exon 12

(Figure 2, lanes 3 and 4, lower band). However, this variant also resulted in reduced

amount of transcripts that retained the exon 12 (Figure 2, lanes 3 and 4, upper band). (Note

the presence of some *BRCA2ΔE12* product, mirroring the situation *in vivo* in humans).

Taken together, these results show that the deletion of exon 12 of BRCA2 does not affect

the viability of mouse ES cells. We then tested the *Brca2*-null ES cells expressing the

BRCA2ΔE12 variant for other known functions of BRCA2. In all experiments, we

compared the results for the *BRCA2ΔE12* variant with those obtained for clones that

expressed p.I2285V and wild-type BRCA2.

 Two clones were then tested extensively for sensitivity to DNA-damaging agents and we

found no significant difference for either the *BRCA2ΔE12* allele or the p.I2285V allele

compared to the control cells for any of the six agents used (Figure 3, panels A-F). As

BRCA2 has a role in the recruitment of RAD51 to the DNA repair sites (Yuan et al., 1999),

we tested the effect of the deletion of exon12 of *BRCA2* and of the p.I2285V allele on

radiation induced RAD51 foci formation. Both the *BRCA2ΔE12* and the p.I2285V allele-

containing cells formed similar a number of foci as the cells expressing wild-type *BRCA2*

(Figure 4). The known role of BRCA2 in homologous recombination and genomic stability

(Khanna and Jackson, 2001) prompted us to test these cells for their efficiency of

homologous recombination and genomic instabilities. For the former, we examined the

efficiency of homologous recombination using a gene-targeting vector as described

previously (Kuznetsov et al., 2008). This targeting vector targets *Gt(ROSA)26Sor*

(abbreviated here as *Rosa26*) locus. We observed similar targeting efficiency in the ES cells

expressing *BRCA2ΔE12*, p.I2285V or wild-type *BRCA2* (Figure 5A). The karyotyping of

these cells also did not reveal any significant increase of chromosomal abnormalities due to

deletion of exon12 (Figure 5B and C). Based on these results we conclude that the deletion

of exon12 of *BRCA2* does not significantly affect any of its known functions. We have

previously demonstrated the utility of the ES cell based assay to detect subtle functional

defects in variants that are predicted to be of low risk (Kuznetsov et al., 2008). However,

the possibility of clinically relevant effects of this allele on breast cancer risk that were not due to the known functions of BRCA2 that we assayed here still remained, so we probed the function of p.12285V from a clinical perspective.

Allelic association study

The loss of exon 12 could have an effect on breast cancer risk that might be insufficiently large to be evident from pedigree analysis, but might still be clinically relevant. To address this question, we assessed the frequency of *BRCA2: c.6853A>G* in on-going series of Ashkenazi Jewish breast cancer cases and ethnically matched controls (total n = 1197). We identified 10 carriers in 475 controls, and 15 carriers in 722 cases, resulting in an odds ratio for the association between this allele and breast cancer of 0.99, 95% confidence interval 0.44-2.21, P = 0.97. When we adjusted for age, the odds ratio became 0.97 (95% CI: 0.41-2.12), P = 0.88. Setting the alpha at 5%, a study of this size has an 80% power to rule out an allele frequency of 0.053 or greater in the cases, when the allele frequency in controls is 0.021.

Discussion

Here, we describe the analysis of *BRCA2: c.6853A>G* (BIC nomenclature, *7081A>G*, p.12285V), a previously reported VUS. We first confirmed that *BRCA2: c.6853A>G* co-occurs in *trans* with other deleterious alleles in *BRCA2*. We also showed that this variant promotes increased exclusion of exon 12 of *BRCA2*, but is not associated with a clinically relevant increase in breast cancer risk. Importantly, we have carried out a comprehensive analysis of the functional consequences of loss of exon 12 of *BRCA2*, using a murine ES

cell system, wherein functional BRCA2 is required to rescue *Brca2-null* murine ES cells

(Kuznetsov et al., 2008).

Inappropriate splicing out of functionally critical exons is associated with disease. Most of

the mutations that affect splicing are those that disrupt splice sites (Hastings and Krainer,

2001). More recently, however, exonic mutations that affect splicing have been identified

(Cartegni et al., 2002). Our results show that *BRCA2: c.6853A>G* is one such variant, and

that a degree of exclusion of exon 12 is seen in alleles without this variant (Figures 1 and

2). It appears that ESEs may be present in exon 12 of BRCA2 (Figure 1A), and it is notable

that I2285, and the surrounding amino acids (KRRGEPS_I_KRNLLNEFDR) are well

conserved through vertebrate evolution. Thus, the question of the significance of loss of

exon 12 is an interesting one. As noted, loss of exon 12 results in a BRCA2 protein that is

shorter than the wild-type form only by 32 amino acids, and this part of BRCA2 does not

contain any known important functional motifs. Nevertheless, *BRCA2ΔE12* has been

reported to be more commonly found in breast cancers than in normal tissues from the

same individuals (Bieche and Lidereau, 1999). In this study of breast cancer and matched

normal tissue from 12 women, the proportion of BRCA2ΔE12 transcripts was higher in the

tumor in four pairs, and in no pairs was the BRCA2ΔE12 higher in the normal tissue than

in the tumor tissue. Similarly, in a study of 7 ovarian cancer cell lines, 5 were found to have

a two- to five-fold increase in BRCA2ΔE12 compared with mean values found in cultures

of normal human ovarian surface epithelium. The same over-expression in tumor cell lines,

rather than normal epithelium, was seen in one of four breast cancer cell lines and one of

four prostate cancer cell lines. In the case of DU-145, the wild-type isoform of BRCA2 was barely detectable, whereas the truncated isoform, BRCA2ΔE12, was strikingly over-expressed (Rauh-Adelmann et al., 2000). Despite these findings, the functional analyses presented here indicate that BRCA2ΔE12 can assume the biological role of full-length BRCA2. Very likely, the increased level of BRCA2ΔE12 in cancer cells is an epiphenomenon rather than a causative mechanism.

We do not think BRCA2ΔE12 represents a clinically important splice variant of BRCA2. Nevertheless, it is important to investigate all potentially deleterious intra-exonic VUS using several approaches. For example, the *BRCA2* variant *c.8393C>G* (p.T2722R) was initially found to result in an out-of-frame fusion of exons 17 and 19 of *BRCA2* (Fackenthal et al., 2002). Later, it was found that the skipping of exon 18 is not complete (erratum, Am J Hum Genet, December 2003) and thus the significance of the original finding was questioned. Subsequently, both genetic (Easton et al., 2007) and functional (Kuznetsov et al., 2008) analyses have confirmed that this variant is deleterious.

Although the allelic association study we have carried out has not ruled out a small effect on risk, 14,199 Ashkenazi Jewish cases and an equal number of controls would be required to rule out an odds ratio of 1.25 (with the same $1-\beta$ and $\acute{\alpha}$ as above). Therefore, it remains possible that the splicing out of exon 12 does result in some subtle perturbations in risk for breast cancer, but at the current time, we cannot detect such an effect, if is exists at all.

The Breast Cancer Information Core (BIC) (http://research.nhgri.nih.gov/bic/) provides easily accessible information on numerous variants, but is dependent upon timely entry of new information about all variants. Based on a new classification system for VUS that was determined by an IARC working group and recently published (Plon et al., 2008), we suggest that this variant should now referred to a class 1 variant (probability of being pathogenic of less than 0.1%). This conclusion is amply supported by other recently-published model-based evidence (Spearman et al., 2008).

Here, we show for the first time, that loss an exon of a tumor suppressor gene has no measurable consequences across a range of activities. The technique described here could be used to comprehensively assess the separate functionality of each exon of *BRCA2*, and the findings raise the question of a possible class of exons that are dispensable for BRCA2 tumor suppressing functions. Indeed, earlier biochemical studies have revealed that heterologous fusion proteins consisting of BRCA2 domains interacting with Rad51 and Dss1 are capable of restoring DNA repair in BRCA2 deficient cells. The BRCA2 domains mediating association with Rad51 and Dss1 are the BRC repeats (central) and OB (oligosaccharide-binding, C-terminal) domains, respectively (Saeki et al., 2006). Notably, the peptide encoded by exon 12 is not involved in either of these domains. Thus the murine ES system used here could be used to investigate which functions, if any, are impaired in cells that lack various combinations of *BRCA2* exons. It is interesting to note that exons 10 and 12 of BRCA2 do not encode functionally critical peptides and are in-frame (i.e. the reading frame will not be altered when these exons are not represented in the transcript). Therefore, missense variants located in these two exons that alter an amino acid and/or

cause skipping of the entire exon are likely neutral. This statement should be tempered by

the possibility that variants could exist that may have unpredictable effect on the stability or

structural integrity of BRCA2 protein. However, such variants are likely to be rare. In

support of this hypothesis, on reviewing the published literature (Easton et al., 2007;

Spearman et al., 2008; Gomez Garcia et al., 2009) and the BIC

http://research.nhgri.nih.gov/bic/, we were unable to identify any known pathogenic

misssense variants in exons 10 or 12 of *BRCA2*. While not conclusive, this observation

supports our contention that exon 12 (and possibly exon 10) is functionally dispensable.

Using the murine ES cell system reported here, it will be interesting to see, if such

redundancy is seen in other multi-exon tumor suppressor genes.

17

References

1. Bieche I, Lidereau R. 1999. Increased level of exon 12 alternatively spliced BRCA2 transcripts in tumor breast tissue compared with normal tissue. Cancer Res 59:2546-2550.

2. Cartegni L, Chew SL, Krainer AR. 2002. Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat Rev Genet 3:285-298.

3. Easton DF, Deffenbaugh AM, Pruss D, Frye C, Wenstrup RJ, len-Brady K, Tavtigian SV, Monteiro AN, Iversen ES, Couch FJ, Goldgar DE. 2007. A systematic genetic assessment of 1,433 sequence variants of unknown clinical significance in the BRCA1 and BRCA2 breast cancer-predisposition genes. Am J Hum Genet 81:873-883.

4. Fackenthal JD, Cartegni L, Krainer AR, Olopade OI. 2002. BRCA2 T2722R is a deleterious allele that causes exon skipping. Am J Hum Genet 71:625-631.

5. Frank TS, Deffenbaugh AM, Reid JE, Hulick M, Ward BE, Lingenfelter B, Gumpper KL, Scholl T, Tavtigian SV, Pruss DR, Critchfield GC. 2002. Clinical characteristics of individuals with germline mutations in BRCA1 and BRCA2: analysis of 10,000 individuals. J Clin Oncol 20:1480-1490.

6. Gomez Garcia EB, Oosterwijk JC, Timmermans M, van Asperen CJ, Hogervorst FB, Hoogerbrugge N, Oldenburg R, Verhoef S, Dommering CJ, Ausems MG, van Os TA, van der Hout AH, Ligtenberg M, van den OA, van der Luijt RB, Wijnen JT, Gille JJ, Lindsey PJ, Devilee P, Blok MJ, Vreeswijk MP. 2009. A method to assess the clinical significance of unclassified variants in the BRCA1 and BRCA2 genes based on cancer family history. Breast Cancer Res 11:R8.

7. Hastings ML, Krainer AR. 2001. Pre-mRNA splicing in the new millennium. Curr Opin Cell Biol 13:302-309.

8. Khanna KK, Jackson SP. 2001. DNA double-strand breaks: signaling, repair and the cancer connection. Nat Genet 27:247-254.

9. Kuznetsov SG, Liu P, Sharan SK. 2008. Mouse embryonic stem cell-based functional assay to evaluate mutations in BRCA2. Nat Med 14:875-881.

10. Liu HX, Cartegni L, Zhang MQ, Krainer AR. 2001. A mechanism for exon skipping caused by nonsense or missense mutations in BRCA1 and other genes. Nat Genet 27:55-58.

11. Mathe E, Olivier M, Kato S, Ishioka C, Hainaut P, Tavtigian SV. 2006. Computational approaches for predicting the biological effect of p53 missense

mutations: a comparison of three sequence analysis based methods. Nucleic Acids Res 34:1317-1325.

12. Mitchell MK, Gregersen PK, Johnson S, Parsons R, Vlahov D. 2004. The New York Cancer Project: rationale, organization, design, and baseline characteristics. J Urban Health 81:301-310.

13. Pal T, Permuth-Wey J, Betts JA, Krischer JP, Fiorica J, Arango H, LaPolla J, Hoffman M, Martino MA, Wakeley K, Wilbanks G, Nicosia S, Cantor A, Sutphen R. 2005. BRCA1 and BRCA2 mutations account for a large proportion of ovarian carcinoma cases. Cancer 104:2807-2816.

14. Pharoah PD, Antoniou AC, Easton DF, Ponder BA. 2008. Polygenes, risk prediction, and targeted prevention of breast cancer. N Engl J Med 358:2796-2803.

15. Plon SE, Eccles DM, Easton D, Foulkes WD, Genuardi M, Greenblatt MS, Hogervorst FB, Hoogerbrugge N, Spurdle AB, Tavtigian SV. 2008. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. Hum Mutat 29:1282-1291.

16. Rauh-Adelmann C, Lau KM, Sabeti N, Long JP, Mok SC, Ho SM. 2000. Altered expression of BRCA1, BRCA2, and a newly identified BRCA2 exon 12 deletion variant in malignant human ovarian, prostate, and breast cancer cell lines. Mol Carcinog 28:236-246.

17. Saeki H, Siaud N, Christ N, Wiegant WW, van Buul PP, Han M, Zdzienicka MZ, Stark JM, Jasin M. 2006. Suppression of the DNA repair defects of BRCA2-deficient cells with heterologous protein fusions. Proc Natl Acad Sci U S A 103:8768-8773.

18. Shih HA, Nathanson KL, Seal S, Collins N, Stratton MR, Rebbeck TR, Weber BL. 2000. BRCA1 and BRCA2 mutations in breast cancer families with multiple primary cancers. Clin Cancer Res 6:4259-4264.

19. Spearman AD, Sweet K, Zhou XP, McLennan J, Couch FJ, Toland AE. 2008. Clinically Applicable Models to Characterize BRCA1 and BRCA2 Variants of Uncertain Significance. J Clin Oncol 26:5393-5400.

20. Tavtigian SV, Byrnes GB, Goldgar DE, Thomas A. 2008. Classification of rare missense substitutions, using risk surfaces, with genetic- and molecular-epidemiology applications. Hum Mutat 29:1342-1354.

21. Tavtigian SV, Deffenbaugh AM, Yin L, Judkins T, Scholl T, Samollow PB, de SD, Zharkikh A, Thomas A. 2006. Comprehensive statistical study of 452 BRCA1

missense substitutions with classification of eight recurrent substitutions as neutral. J Med Genet 43:295-305.

22. Warming S, Costantino N, Court DL, Jenkins NA, Copeland NG. 2005. Simple and highly efficient BAC recombineering using galK selection. Nucleic Acids Res 33:e36.

23. Yang Y, Sharan SK. 2003. A simple two-step, 'hit and fix' method to generate subtle mutations in BACs using short denatured PCR fragments. Nucleic Acids Res 31:e80.

24. Yuan SSF, Lee SY, Chen G, Song MH, Tomlinson GE, Lee EYHP. 1999. BRCA2 is required for ionizing radiation-induced assembly of rad51 complex in vivo. Cancer Res 59:3547-3551.

Figure Legends

Figure 1: Splicing at *BRCA2: c.6853A>G* in humans

A) Potential exonic splicing enhancer (ESEs) disrupted by BRCA2: c.6853A>G

(p.I2285V)

RESCUE-ESE (http://genes.mit.edu/burgelab/-rescue-ese/) identified putative ESEs at the

region where the *BRCA2: c.6853A>G* variant is located. Notably, the variant "G" allele

disrupts three of overlapping putative ESEs.

B) Splicing of BRCA2 exon 12 in humans carrying the c.6853A>G allele

The amplicon of exon 12-13 was sequenced using a reverse primer located at exon 13. The

position of c.6853 in the sequencing chromatogram is indicated by a dashed box. The wild-

type A allele (shown here as T on the reverse strand) is the major signal in the three

c.6853A>G carriers (I2285V-a, I2285V-b and I2285V-c) and a non-carrier control

(control). The cell line DBV carries double variants of 6174delT and c.6853A>G. The

result shown for DBV suggests that there is no technical allele bias in the PCR. WLT is a

wild-type control line carrying neither variant.

C) Relative quantification of exon 12 skipping in WLT and DBV LCLs

Three amplicons were produced: E12-13 with exon 12 inclusion, ΔE12 with exon 12

skipping and E11 as a control of endogenous expression level. The primers are listed in the

methods section. **C**: Semiquantitative comparison of exon 12 skipping in WTL and DBV:

normalization was done by considering E11 expression as 1.00. The relative expression

level of E12 intact vs. ΔE12 is shown as a percentage bar. Blue bars: exon 12 skipping

transcripts, black bars: exon 12 intact transcripts. The error bar indicates the variance seen

following four measurements.

Figure 2: *RT-PCR demonstrating 96-bp deletion in p. I2285V and BRCA2ΔE12 mutants.*

Ethidium Bromide stained gel showing RT-PCR results of two mouse embryonic stem cell

clones generated independently for each variant. WT: wild-type, negative control: water

control, FL: fragment of a full length product, ΔE12: transcript with deletion of exon 12.

Numbers below the gel picture are lane numbers. Marker is 1 kb plus ladder (Invitrogen).

Figure 3: *Sensitivity of ES cells expressing mutant and wild-type BRCA2 to different*

DNA-damaging agents.

Survival of ES cells expressing p.I2285V, BRCA2ΔE12 mutants compared to ES cells

expressing wild type BRCA2 after exposure to different DNA-damaging agents: Methyl-

N'-nitro-N-Nitrosoguanidine (MNNG, A), mitomycin C (MMC, B), cisplatin (C), methyl-

methanesulfonate (MMS, D), ionizing radiation (IR, E) and ultraviolet light (UV, F). Drug

sensitivity was expressed as a percentage of surviving cells compared to untreated cultures.

Two independently generated clones for each variant were tested and their average values

are shown here. Error bars indicate standard deviations. *P* values were equal 0.19693

(I2285V); 0.16019 (BRCA2ΔE12) for 5 μM MNNG, 0.09178 (I2285V); 0.08212

(BRCA2ΔE12) for 10 ng/ml MMC, 0.5494 (I2285V); 0.4110 (BRCA2ΔE12) for 0.4 μM

cisplatin, 0.11846 (I2285V); 0.74637 (BRCA2ΔE12) for 15 μg/ml MMS, 0.1131 (I2285V);

0.14066 (BRCA2ΔE12) for 200 Rad IR and 0.3964 (I2285V); 0.4238 (BRCA2ΔE12) for

10 J/m^2 UV.

Figure 4: Radiation-induced RAD51 foci formation in mouse ES cells expressing BRCA2 transgenes.

A) RAD51 foci formation before (-IR) and after (+IR) ionizing radiation. RAD51 foci are shown in green, γ-H2AX foci marking sites of DNA damage are shown in red, nuclei are stained with DAPI (blue). The right-most panel shows the merged picture. Two independent clones showed similar results and only one of them is shown in each case.

B) Quantification of RAD51 and γ-H2AX foci before (-IR) and after (+IR) ionizing radiation. 30 nuclei were counted in each case and their mean values are shown. Error bars indicate means ± s.d.

Figure 5: Effect of BRCA2 variants on homologous recombination and genomic integrity in mouse ES cells.

a) *Homologous recombination efficiency* shown by gene targeting at *Rosa26* locus. Numbers above the bars indicate the actual number of colonies undergoing homologous recombination compared to total number of colonies tested.

b) *Karyotype analysis of p.I2285V, BRCA2ΔE12 mutants compared to WT control cells.* Representative metaphases of the mutants and WT control cells are shown. Ten metaphases were counted in each case and scored for different chromosomal abnormalities. **c)** The identified chromosomal aberrations are tabulated.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

A

GAAACT
AGAAAC
AAGAAA
AAAGAA
AAAAGA
CAAAAG
ATCAAA
AATCAA
CAATCA
GAGAACCCTCAATCAAAAGAAACTTA
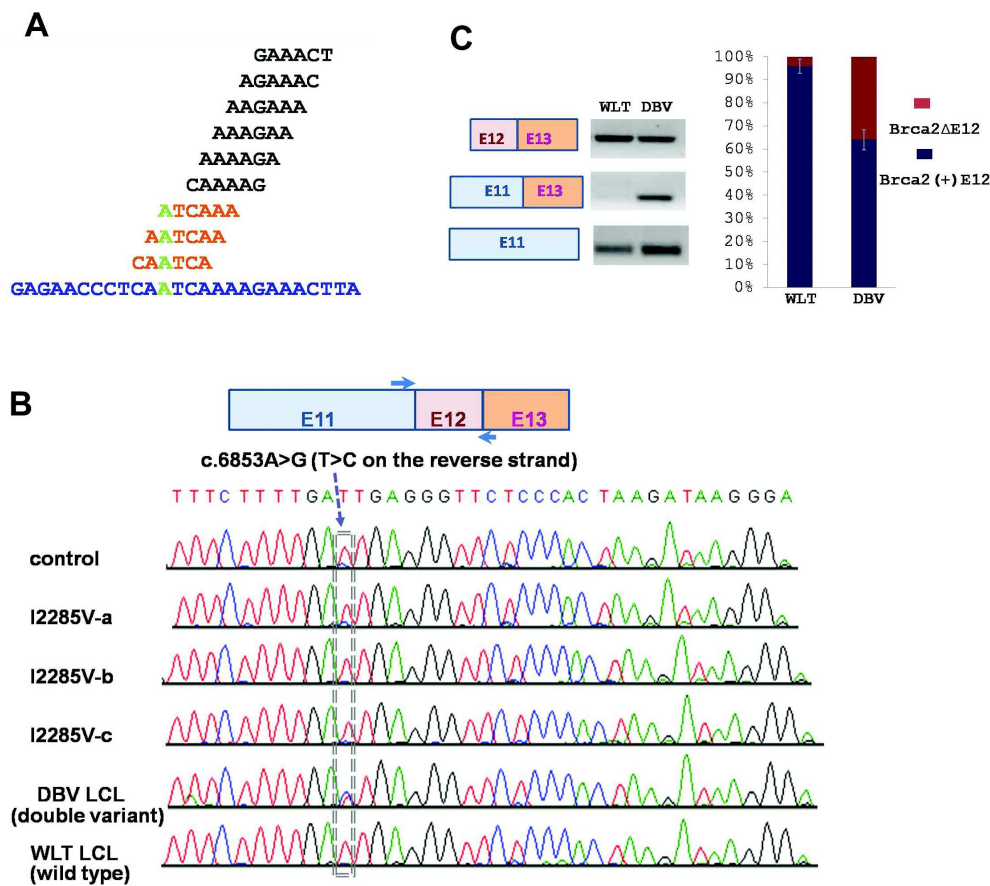
C

B

c.6853A>G (T>C on the reverse strand)

Figure 1: Splicing at BRCA2: c.6853A>G in humans
A) Potential exonic splicing enhancer (ESEs) disrupted by BRCA2: c.6853A>G (p.I2285V)
B) Splicing of BRCA2 exon 12 in humans carrying the c.6853A>G allele
C) Relative quantification of exon 12 skipping in WLT and DBV LCLs

190x173mm (600 x 600 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
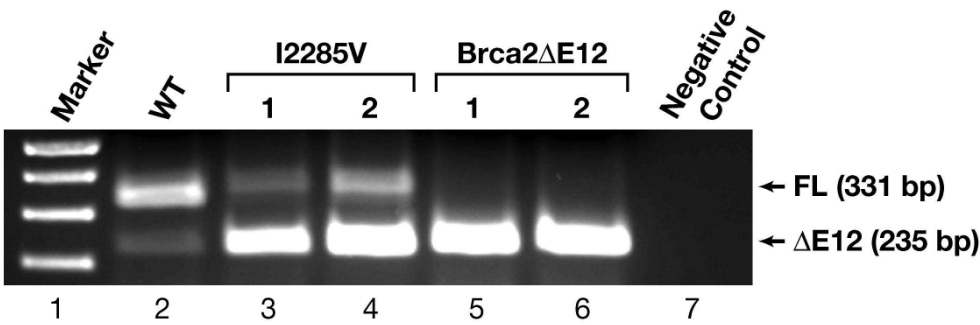52
53
54
55
56
57
58
59
60

Figure 2: RT-PCR demonstrating 96-bp deletion in p. I2285V and BRCA2ΔE12 mutants.
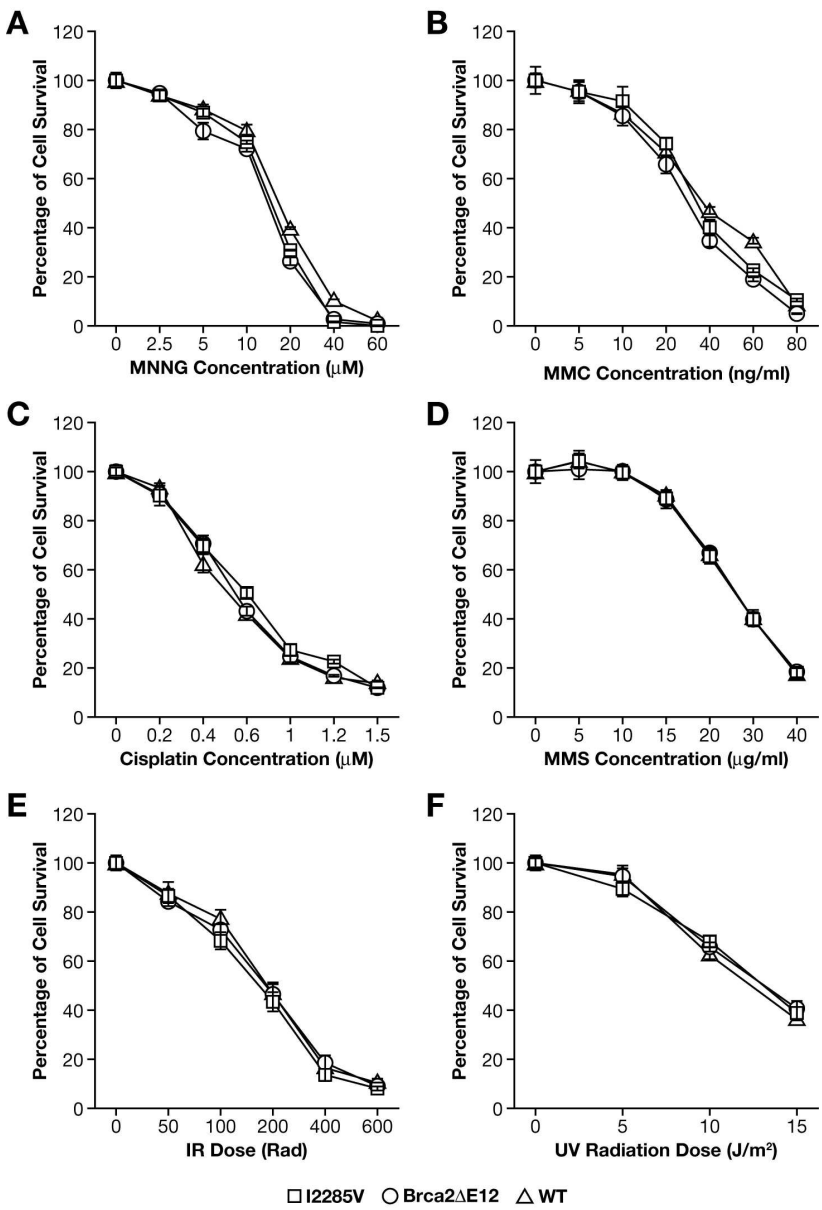96x31mm (600 x 600 DPI)

Figure 3: Sensitivity of ES cells expressing mutant and wild-type BRCA2 to different DNA-damaging agents.
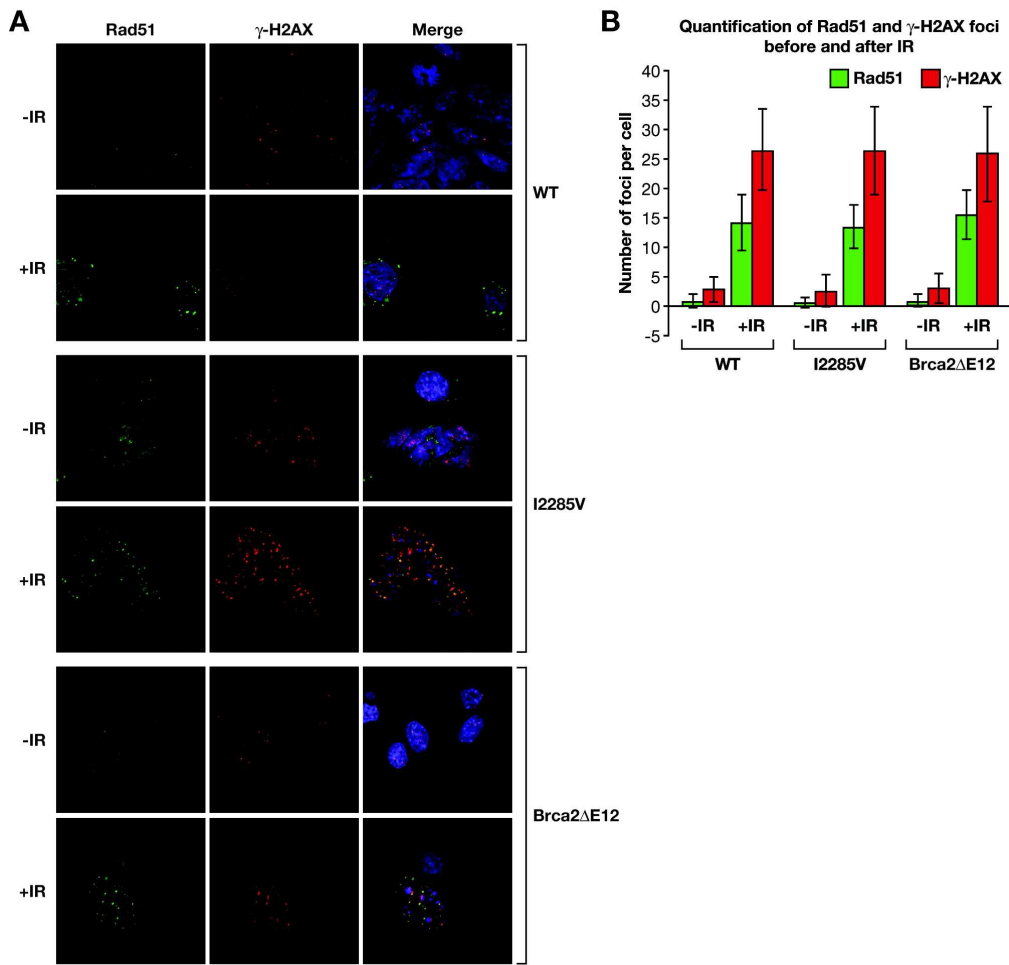123x182mm (600 x 600 DPI)

Figure 4: Radiation-induced RAD51 foci formation in mouse ES cells expressing BRCA2 transgenes.
A) RAD51 foci formation before (-IR) and after (+IR) ionizing radiation.
B) Quantification of RAD51 and γ-H2AX foci before (-IR) and after (+IR) ionizing
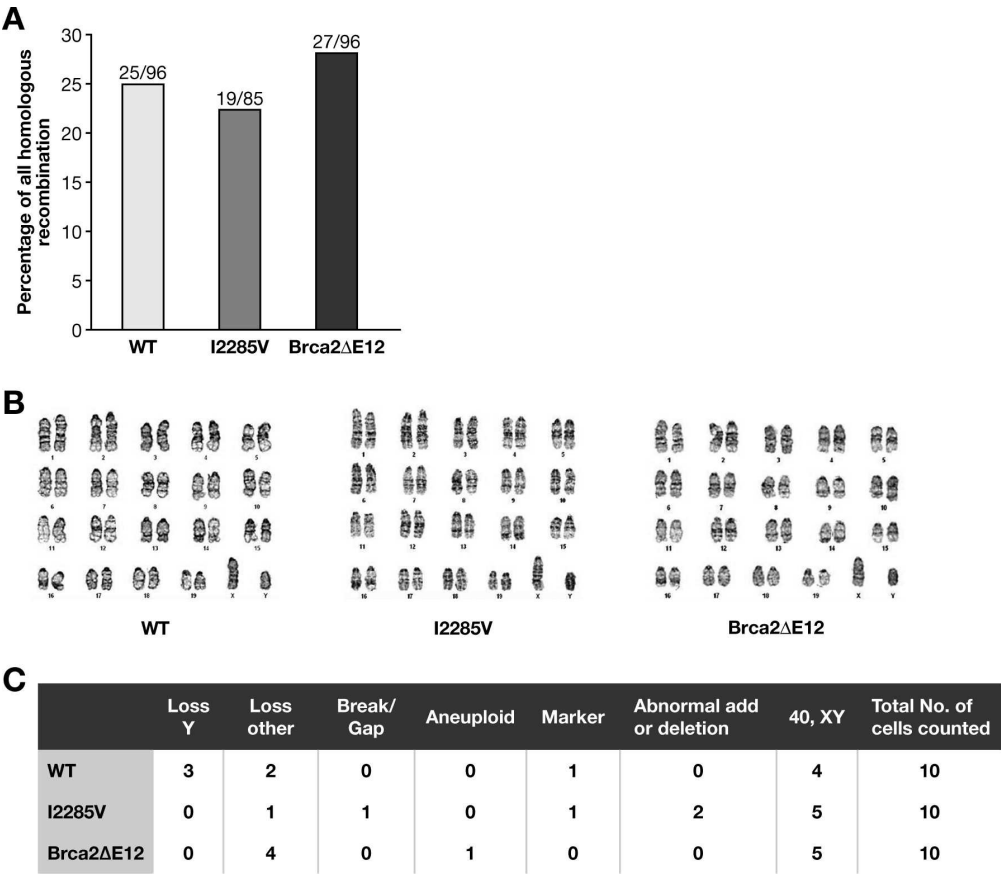
169x161mm (600 x 600 DPI)

A

B

C

	Loss Y	Loss other	Break/ Gap	Aneuploid	Marker	Abnormal add or deletion	40, XY	Total No. of cells counted
WT	3	2	0	0	1	0	4	10
I2285V	0	1	1	0	1	2	5	10
Brca2ΔE12	0	4	0	1	0	0	5	10

Figure 5: Effect of BRCA2 variants on homologous recombination and genomic integrity in mouse ES cells.
a) Homologous recombination efficiency
b)   Karyotype analysis of p.I2285V, BRCA2ΔE12 mutants compared to WT control cells.

151x131mm (600 x 600 DPI)

1
2
3
4
5
6 SUPPLEMENTARY FIGURE LEGENDS
7
8
9
10 SUPPLEMENTARY FIGURE S1
11
12
13 *Pedigrees of four families carrying BRCA2: p. I2285V*
14
15
16 FIGURE 1A
17
18 INDIVIDUAL V:2 PRESENTED TO THE SMBD-JEWISH GENERAL HOSPITAL, MONTREAL, QUEBEC, REQUES
19
20 GENETIC TESTING FOR A BRCA2 VARIANT THAT HE SAID HAS BEEN IDENTIFIED IN HIS FAMILY. HE
21
22 ALONG GENETIC TESTING RESULTS WHICH ALLOWED US TO IDENTIFY THE VARIANT AS I2285V. TH
23
24 ASHKENAZI JEWISH ORIGINS AND WAS ORIGINALLY TESTED IN ANOTHER CANADIAN PROVINCE. II
25
26 V:3 AND IV:2 HAD ALREADY BEEN FOUND TO CARRY THIS VARIANT. WE DECLINED TO OFFER CLIN
27
28 V:1 BUT ASKED IV: 2 TO COME IN FOR MORE DETAILED STUDIES. HAVING CONFIRMED THE PRESENC
29
30 VARIANT, SAMPLES FROM THIS INDIVIDUAL WERE USED TO FURTHER CHARACTERIZE THE VARIA
31
32 FIGURE 1B
33
34 INDIVIDUAL III:1 PRESENTED FOR GENETIC TESTING AT THE MCGILL UNIVERSITY HEALTH CENTRE, M
35
36 QUEBEC, ON ACCOUNT OF HER DIAGNOSIS OF SEROUS PAPILLARY OVARIAN CANCER DIAGNOSED A
37
38 TOGETHER WITH HER ASHKENAZI JEWISH ANCESTRY. A THREE GENERATION PEDIGREE WAS OBTAIN
39
40 PATHOLOGY RECORDS CONFIRMED THE DIAGNOSIS IN III:1. PATHOLOGY RECORDS WERE NOT AVAIL
41
42 REMAINING RELATIVES. GENETIC TESTING FOR THE THREE ASHKENAZI JEWISH FOUNDER MUTATIC
43
44 OFFERED TO INDIVIDUAL III:1 AND REVEALED THE PRESENCE OF A BRCA2:6174DELT MUTATION. FC
45
46 IDENTIFICATION OF A BRCA2 MUTATION IN THE FAMILY, INDIVIDUAL III:2 PRESENTED FOR GENET.
47
48 SUBSEQUENT TESTING FOR THE THREE FOUNDER MUTATIONS ALSO IDENTIFIED INDIVIDUAL III:2 A
49
50 THE BRCA2 6174DELT MUTATION. AT THE TIME OF TESTING, III:2 ALSO PROVIDED A BLOOD SAMPL
51
52 RESEARCH PURPOSES. IN THE COURSE OF OUR EFFORTS TO IDENTIFY A COMPOUND HETEROZYGOTE, SI

C0NF0RMATI0N ANALYSIS REVEALED THE PRESENCE 0F THE BRCA2 I2285V VARIANT. FURTHER T

THE LYMPH0CYTE DNA FR0M INDIVIDUAL III:1 ALS0 REVEALED THE EXISTENCE 0F THE BRCA2 I2285

VARIANT. N0 0THER FAMILY MEMBERS WERE AVAILABLE F0R TESTING.

<u>FIGURES 1C AND 1D</u>

B0TH FAMILIES ALL HAVE ASHKENAZI JEWISH ANCESTRY AND ALL HAD FULL SEQUENCING 0F BR0

MYRIAD GENETICS LAB0RAT0RIES (SALT LAKE CITY, UT). IN FIGURE 1C, INDIVIDUAL III.1 WAS SE

INITIAL VISIT AND TESTED IN 2004, WHEN THE BRCA2: I2285V VARIANT WAS IDENTIFIED. IN FIGU

III.1 WAS TESTED ELSEWHERE IN 2007 AND THE BRCA2: I2285V VARIANT WAS IDENTIFIED. LATER

YEAR, THE PR0BAND CAME T0 MEM0RIAL SL0AN KETTERING CANCER CENTER F0R AN INITIAL VIS

THIS RESULT.

NB IN ALL FIGURES THE PR0BAND IS INDICATED WITH AN ARR0W.

SUPPLEMENTARY FIGURE S2

A D0UBLE VARIANT LYMPH0BLAST0ID CELL LINE (LCL) REFERRED T0 AS DBV WAS ESTABLISHED

INDIVIDUAL CARRYING B0TH BRCA2: NM_000059.3:C.5946DELT (P.SER1982ARGFSX20, THE F0UND

MUTATI0N KN0WN AS 6174DELT, ACC0RDING T0 BIC) AND BRCA2: NM_000059.3: C.6853A>G

(P.I2285V, 7081A>G ACC0RDING T0 BIC). THE T0P TW0 R0WS 0F SEQUENCING ELECTR0GRAMS SH0

THE PRESENCE 0F THE D0UBLE VARIANTS IN THE GEN0MIC DNA 0F THIS CELL LINE. F0R GEN0MIC

SEQUENCING , PRIMER PAIRS 0F F0RWARD 5'-AGGTTGTTACGAGGCATTGG-3' WITH REVERSE 5'-

TCCAGAGAA AG CAGATG AATTT-3'; AND F0RWARD 5'-

TCTGTGGTATCTGGTAGCATCTG-3' WITH REVERSE 5'-GCACA GTGGCTCATGTCTGT-3' ,

WERE USED T0 AMPLIFY THE DNA REGI0NS 0F C.5946DELT AND C.6853A>G, RESPECTIVELY. F0RWA

PRIMERS WERE USED AS SEQUENCING PRIMERS. ALLELE-SPECIFIC PCR WITH CDNA WAS PERF0RM

F0RWARD PRIMERS DIFFERENTIATING THE C.5946T WILD-TYPE ALLELE AND C.5946DELT ALLELE: 5'

GGGATTTTTAGCACAGCAAGG -3' (F0R C.5946DELT); 5'-

TTGTGGGATTTTTAGCACAGCAAGT-3' (F0R THE WILD TYPE ALLELE AT C.5946T), THE REVERSE

PRIMER 0F 5'-GTGCGAAAGGGTACACAGGT-3' WAS USED F0R AMPLIFYING B0TH ALLELES.

C.5946T WILD TYPE ALLELE-SPECIFIC PCR WAS PERF0RMED WITH CDNA FR0M A DBV LCL AND A

TYPE LCL, B0TH EX0N 12 INCLUSI0N (0F 1085BP) AND SKIPPING (0F 989BP) TRANSCRIPTS WERE DE

THE SCHEMA F0R THE ALLELE-SPECIFIC PCR REACTI0NS IS SH0WN IN THE NEXT TW0 FIGURES (T0

C.5946DELT ALLELE; B0TT0M C.6853A>G ALLELE).

THE NEXT PANEL SH0WS THAT THE DEGREE 0F EX0N 12 SKIPPING FR0M THE WILD-TYPE ALLELE (C

HIGHER IN THE DBV LINE THAN IN WILD-TYPE INDIVIDUALS. IN THE FINAL R0W, CDNA FR0M DB

WAS USED F0R ALLELE-SPECIFIC PCR, BY USING C.5946T (N0RMAL) AND C.5946DELT PRIMER,

RESPECTIVELY. SEQUENCING THE EX0N 12 INCLUSIVE TRANSCRIPT FR0M THE ALLELE 0F C.5946DE

THE REVERSE PRIMER SH0WED "A" (T 0N THE REVERSE STRAND) AT THE P0SITI0N 0F 6853, ILLUSTR

C.5946DELT AND C.6853A>G ARE IN *trans*

1
2
3
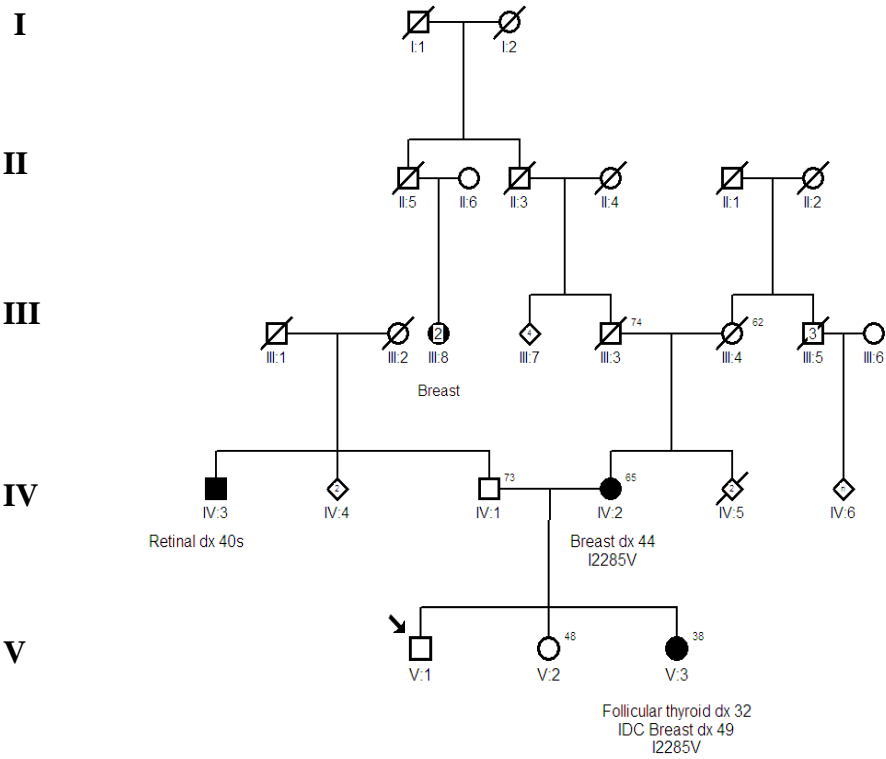4 **Figure S1: Pedigrees of four independent probands carrying the BRCA2: p.I2285V allele**
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32 **Figure 1a**
33

34
35
36
37
38
39
40
41
42
43
44
45
46
47
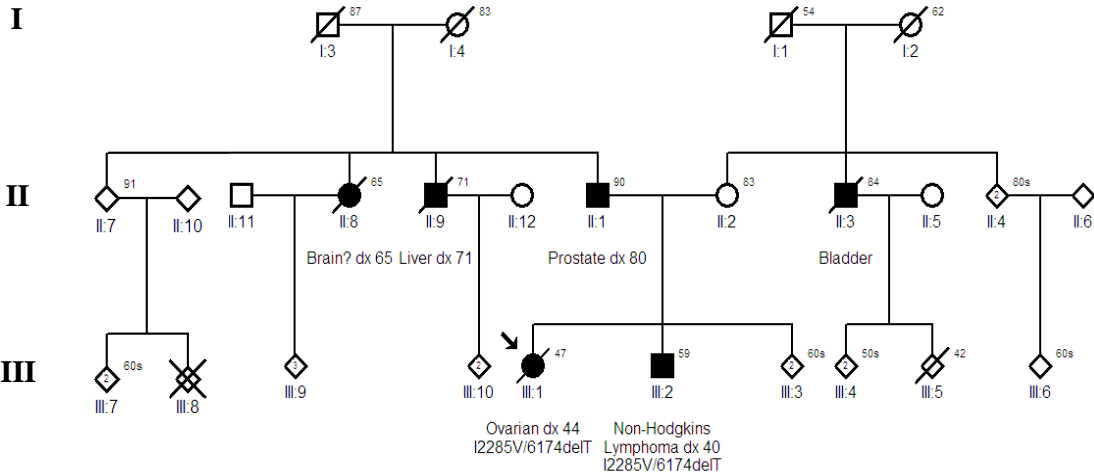48
49
50
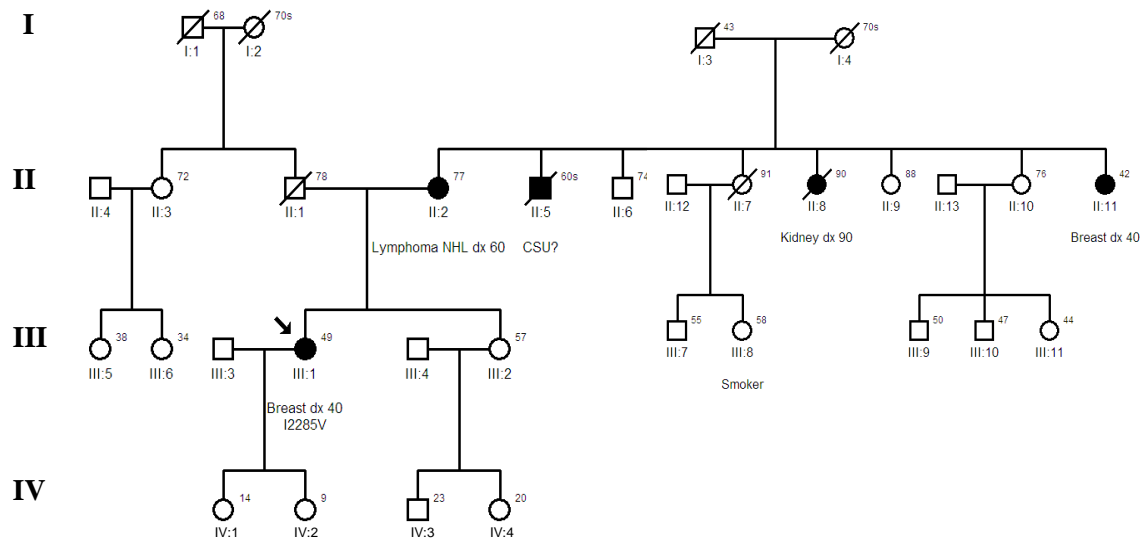51
52
53
54
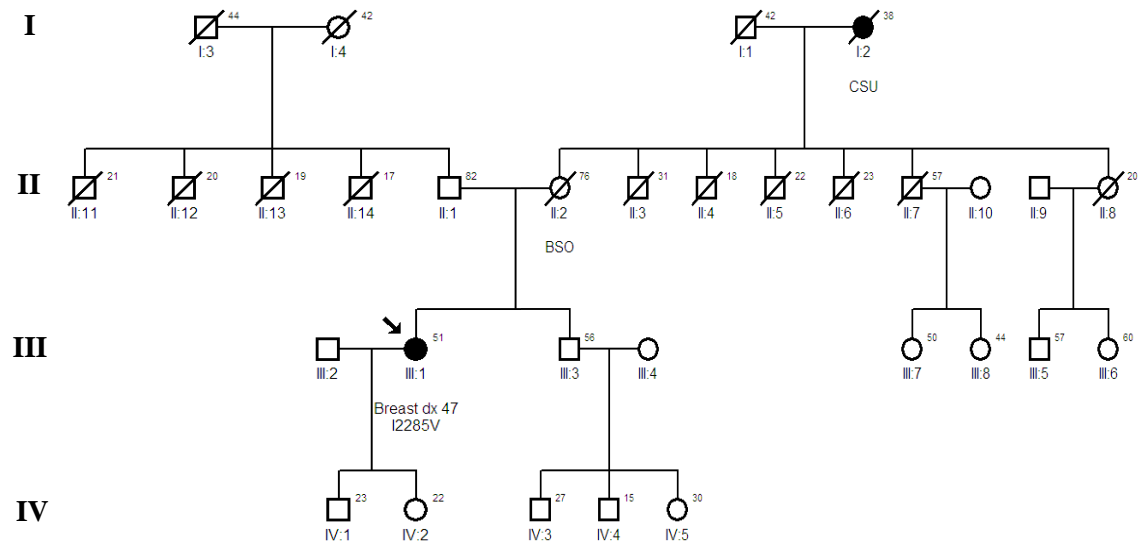55 **Figure 1b**
56
57
58
59
60

Figure 1c

Figure 1d

FigureS2: BRCA2 founder mutation c.5946delT and variant c.6853A>G are *in-trans*